MTH211 - Assignment 2

Group -G

2023-05-04

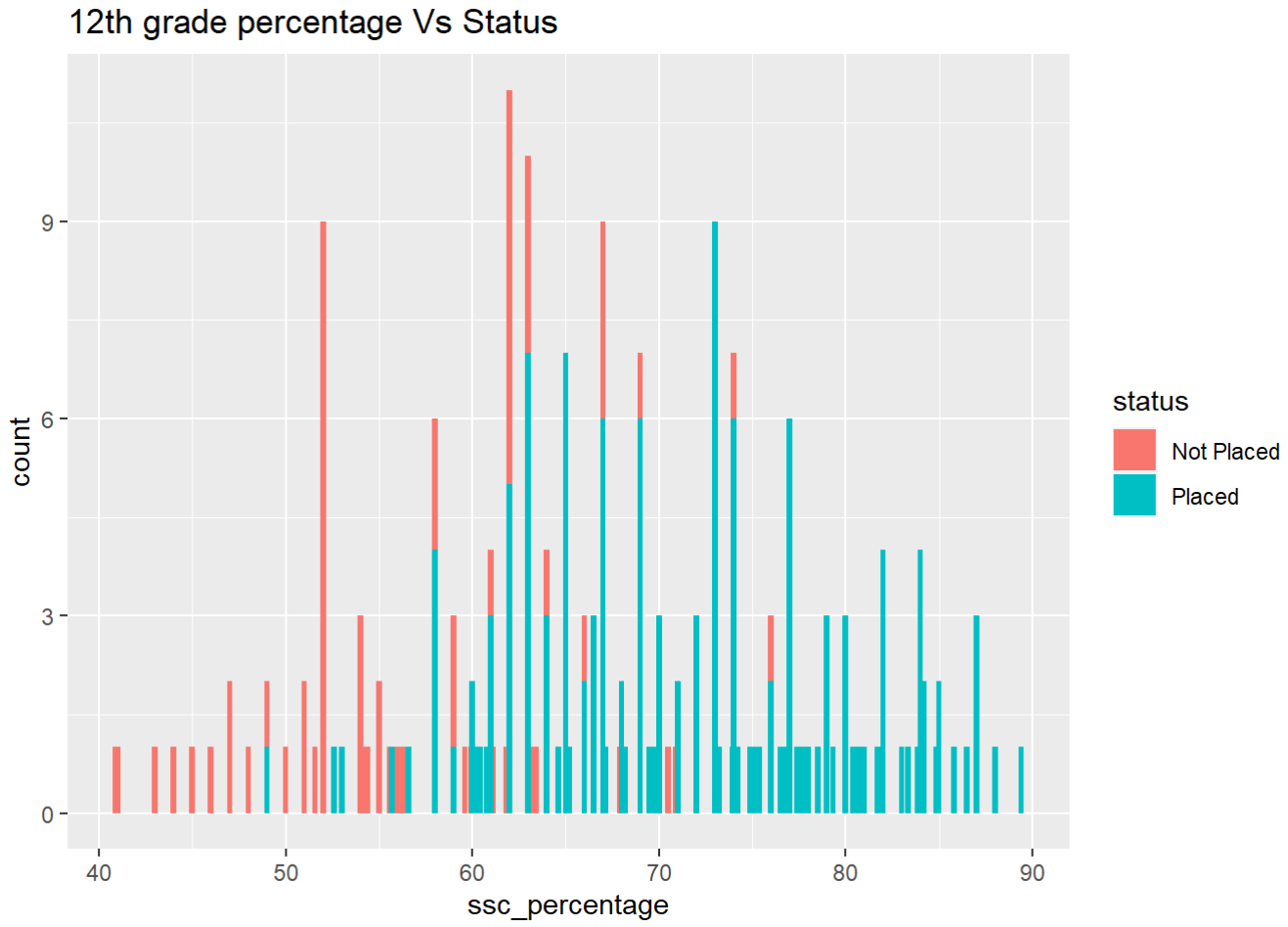
It seems that the percentage in 12th exam has a significant association with placement. Test, with appropriate assumptions, if the distribution of percentage in 12th grade differs significantly between placed people and not placed people.

We have used dataset for job placement present at Kaggle - “Data-set” (https://www.kaggle.com/datasets/ahsan81/job-placement-dataset?resource=download)

```
job_dat <- read.csv("Job_Placement_Data.csv")
placed <- job_dat$hsc_percentage[job_dat$status == "Placed"]
not_placed <- job_dat$hsc_percentage[job_dat$status == "Not Placed"]
```

Here **placed** variable is people’s 12th grade percentage who are placed , **not_placed** variable is people’s 12th grade percentage who are not placed.

```
library(ggplot2)
ggplot(job_dat,aes(ssc_percentage ,fill = status)) + geom_bar(  width = .3)  + labs(title="12th grade percentage Vs Status")
```



We have assumed that the distribution of 12th grade percentage with placed as *Normal distribution* with parameters (μ_1, σ_1^2) and the another distribution of 12th grade percentage with not_placed as *Normal distribution* with parameters (μ_2, σ_2^2)

Hypothesis:
H_0: $\sigma_1^2 = \sigma_2^2$ ***VS*** H_1: $\sigma_1^2 \neq \sigma_2^2$

We are using F-test if our null hypothesis is true or not. For that we have used “var.test” function.

```
var.test(placed, not_placed,"two.sided",ratio = 1)
```

```
##
##  F test to compare two variances
##
## data:  placed and not_placed
## F = 0.8855, num df = 147, denom df = 66, p-value = 0.542
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.5760373 1.3172148
## sample estimates:
## ratio of variances
##          0.8855017
```

As p-value (= 0.542) > 0.05 therefore, we will accept the null hypothesis i.e, $\sigma_1^2 = \sigma_2^2$.

Hypothesis:
H_0: $\mu_1 = \mu_2$ ***VS*** H_1: $\mu_1 \neq \mu_2$ **given that** $\sigma_1^2 = \sigma_2^2 = \sigma^2$

To check this hypothesis we are using “t.test” function.

```
t.test(placed, not_placed,"two.sided",var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  placed and not_placed
## t = 8.2307, df = 213, p-value = 1.849e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   8.76949 14.29257
## sample estimates:
## mean of x mean of y
##  69.92655  58.39552
```

Here, p-value(=1.849e-14) < 0.05 therefore, we will reject our null hypothesis i.e, $\mu_1 \neq \mu_2$

Hence, we can conclude that distributions of placed and not placed people with respect to there 12th grade percentage are not equal.